Human Molecular Genetics, 1998, Vol. 7, No. 3 525–532

Analysis of the dynamic mutation in the SCA7 gene shows marked parental effects on CAG repeat transmission

Launce G. Gouw[1], Marco A. Castañeda[5], Catherine K. McKenna[2,4], Kathleen B. Digre[2,3], Stefan M. Pulst[6], Susan Perlman[6], M. S. Lee[7], Christopher Gomez[8], Kenneth Fischbeck[9], Deborah Gagnon[10], Elsdon Storey[11], Thomas Bird[12], F. Raul Jeri[5] and Louis J. Ptácek[1,2,4,*]

Departments of [1]Human Genetics, [2]Neurology and [3]Ophthalmology and [4]Howard Hughes Medical Institute, University of Utah, Salt Lake City, UT 84112, USA, [5]Neurology Service, Dos de Mayo National Hospital, Lima, Peru, [6]Division of Neurology, Cedars-Sinai Medical Center, UCLA School of Medicine, Los Angeles, CA 90048, USA, [7]Department of Neurology, Youngdong Severance Hospital, Yonsei University College of Medicine, Seoul, Korea, [8]Department of Neurology, University of Minnesota, Minneapolis, MN 55455, USA, [9]Department of Neurology, University of Pennsylvania, Philadephia, PA 19104, USA, [10]Department of Pediatrics, University of Illinois, Chicago, IL 60612, USA, [11]Murdoch Institute, Royal Children's Hospital, Parkville, Victoria 3052, Australia and [12]Department of Neurology, VA and Medical Center, University of Washington Medical School, Seattle, WA 98101, USA

Received October 24, 1997; Revised and Accepted December 12, 1997

The gene for spinocerebellar ataxia 7 (SCA7) includes a transcribed, translated CAG tract that is expanded in SCA7 patients. We have determined expansions in 73 individuals from 17 SCA7 kindreds and compared them with repeat lengths of 180 unaffected individuals. Subjects with abnormal expansions comprise 59 clinically affected individuals and 14 at-risk currently unaffected individuals predicted to carry the mutation by haplotype analysis. For expanded alleles, CAG repeat length correlates with disease progression and severity and correlates inversely with age of onset. Increased repeat lengths are seen in generational transmission of the disease allele, consistent with the pattern of clinical anticipation seen in these kindreds. Repeat lengths in expanded alleles show somatic mosaicism in leukocyte DNA, suggesting that these alleles are unstable within individuals as well as between generations. Although dynamic repeat expansions from paternal transmissions are greater than those from maternal transmissions, maternal transmission of disease is more common, suggesting germline or embryonic effects of the repeat expansion.

INTRODUCTION

The autosomal dominant spinocerebellar ataxias (SCAs) are a heterogeneous group of hereditary diseases manifested by degeneration of the cerebellum (primarily cortex) with associated variable pathology in other neural structures, including the inferior olivary nucleus, basal ganglia, spinal cord, retina and peripheral nerves. Clinical manifestations of progressive ataxia, dysarthria and dysmetria may be accompanied by an associated constellation of associated deficits including ophthalmoplegia, visual loss, dementia, sensory deficits and pyramidal and/or extrapyramidal signs, depending on the underlying genetic defect. Previously defined on the basis of clinical (1) and pathological (2) critera, this complex group of diseases has been delineated by the discovery of distinct disease-causing genetic loci (3–8). More recently, successful cloning and determination of several disease-causing mutations have been accomplished for many of these diseases (9–13). We (14) have described families with SCA distinguished by associated pigmentary macular dystrophy and retinal degeneration leading to blindness, which we designated SCA7 (15). Previous nomenclature for SCA7 includes autosomal dominant cerebellar ataxia (ADCA) type II (1) and olivopontocerebellar ataxia (OPCA) type III (2). We (15) and others (16,17) previously have mapped the disease-causing gene to the short arm of chromosome 3 (3p12–13). Following this linkage information, SCA7 recently has been characterized (18).

The striking anticipation seen in this disease (14) mirrors that seen in other SCAs including SCA1, SCA2 and SCA3/Machado–Joseph disease (MJD), as well as other neurodegenerative diseases such as Huntington's disease (HD), dentatorubralpalli-doluysian atrophy/Haw River/Smith's disease (DRPLA) and spinal bulbar muscular atrophy/Kennedy's disease (SBMA). The underlying mutation in these neuropathologies is a transcribed, translated CAG repeat coding for a polyglutamine tract that is expanded in affected individuals [for review, see (19,20)]. As predicted by clinical (14), cellular (21) and molecular (22) studies, this appears to be the case in SCA7 as well. We have

*To whom correspondence should be addressed. Tel: +1 801 581 3993; Fax: +1 801 585 5597; Email: ptacek@genetics.utah.edu

investigated the length of the polyglutamine tract in affected individuals from several large previously described kindreds (15) as well as from other large and several smaller kindreds diagnosed as carrying the phenotype attributed to SCA7, and we have compared repeat length with a large sample of unrelated unaffected individuals.

RESULTS

Repeat length determination in affected and unaffected individuals

Repeat length distribution among three population groups studied is shown in Figure 1a. The distribution of unexpanded alleles is similar for CEPH, unaffected and affected SCA7 kindred individuals (Fig. 1b), with $(CAG)_{10}$ by far the most common allele (71.5% of all unexpanded alleles). CAG repeat lengths ranged from six to 17 in 427 unexpanded chromosomes, with a mean of 10.5 ± 1.1 and median of 10 repeats. Unexpanded alleles showed no mutation, and genotype distribution in this group is consistent with Hardy–Weinberg equilibrium. Unrelated unaffected individuals taken from the CEPH panel had 34.3% observed repeat length heterozygosity (36/105 individuals); 36.6% by the HET equation (210 chromosomes). Unaffected SCA7 kindred subjects incorporating both marry-ins and asymptomatic individuals not carrying the disease genotype showed greater heterozygosity (54.3% observed, 57.1% by HET, $n = 144$ chromosomes). This may be due to the fact that the CEPH reference panel has a preponderance of individuals of European origin, whereas many of the SCA7 kindreds are drawn from a broader geographic region (see Materials and Methods).

Seventy three expanded alleles (>17 repeats) from 59 patients and 14 asymptomatic at-risk individuals carrying the disease genotype ranged from 34 to 103 CAG repeats (Fig. 1c). The mean repeat length among these alleles was 49.4 ± 11.0 with a median of 47 repeats. Analysis of expanded repeats showed somatic variation in DNA derived from peripheral blood lymphocytes (Fig. 2), with expanded repeat lengths varying from ±1 to ±3 repeats from the most prominent band per sample; such heterogeneity was not seen in unexpanded alleles. Although the possibility of PCR infidelity for larger expansions cannot be ruled out, the use of polymerase with proofreading/exonuclease activity makes this less likely.

Repeat length correlation with age of onset, disease severity and generation

As is seen in other trinucleotide repeat expansion diseases with anticipation, there is an inverse correlation between repeat length and age of onset. Age of onset in 59 clinically affected individuals ranged from 3.5 to 60 years. Mean age of onset was 29.9, with a standard deviation of 14.4 years. Correlation between SCA7 age of onset and CAG repeat length using quadratic regression analysis gave a Pearson coefficient of $r = -0.80$ ($P<0.0001$) (Fig. 3). Unlike the situation described in SCA3/MJD (23), the sex of the affected individual does not appear to affect this correlation greatly. When disease was categorized by number of decades between clinically detectable and debilitating disease, correlation was noted ($r = 0.62$, $P<0.0001$) by simple regression analysis.

Consistent with anticipation previously observed in many of the SCA7 kindreds, as age of onset decreased, repeat length increased in each generation. Kindreds where three generations of SCA7 patients were studied had an average age of onset of 48.3 ± 6.4 for generation I, 31.6 ± 10.7 for II and 15.6 ± 9.0 for III. Inverse correlation is seen in the expanded allele data, with a mean CAG repeat length of 42.1 ± 4.3 (range 34–47, $n = 7$) for the first generation, 47.1 ± 5.0 (39–57, $n = 16$) for the second and 54.7 ± 16.0 (38–103, $n = 23$) for the third generation. Smaller kindreds with a two generation history of SCA7 showed an average age of onset of 42.4 ± 10.0 years with a mean repeat length of 45.5 ± 5.5 repeats (40–58, $n = 11$) for generation I and 19.2 ± 5.7 years and 49.6 ± 8.5 repeats (39–66, $n = 16$) for generation II.

Effects of transmitting parent on repeat length expansion

Parent–child transmission of the disease allele showed expansion of the CAG repeat in 32 of 44 cases. The repeat length differential between generations ranged from –13 to +62, with a mean total expansion of 7 ± 3.5 (mode = 3) repeats per transmission. A significant ($P<0.001$) difference between paternal and maternal transmission of the dynamic mutation exists in our kindreds as has been observed in other neurodegenerative trinucleotide repeat expansion diseases (24) [reviewed in (19)]. Whereas 33 maternal transmissions produced anywhere from –13 to +17 CAG repeats in offspring averaging 3.8 ± 3 repeat expansions, 11 paternal transmissions ranged from –6 to +62 repeats in offspring, averaging 16.6 ± 11 repeats (Fig. 4). Nine of the 10 as well as the two largest decreases in repeat length (–13, –8) were due to maternal transmission, and the four largest increases (+24, +24, +38, +62) were due to paternal transmission. However, the average degree of anticipation is similar betweeen the sexes, with a mean age of onset differential from paternal transmissions of 20.2 ± 10.5 years and of 20.2 ± 9.5 years from maternal transmissions.

In our kindreds, it is notable that whereas repeat length expansion is more dramatic in paternal transmissions, the majority of actual transmissions of expanded alleles are maternal in a 3:1 ratio (33 maternal versus 11 paternal, $P<0.001$). If generational disease histories are taken into account (where DNA, and therefore repeat length data was unavailable), the female:male ratio of the disease-transmitting parent increases to 4:1 (65 female versus 16 male disease-transmitting parents). Furthermore, this bias in maternal disease transmission was seen in all large kindreds and in all of the smaller kindreds except two. It should be noted that the total number of females in our study was slightly greater than males, but this difference was not statistically significant ($P>0.05$).

Whereas an expanded (>17 repeats) parental allele was necessary for transmission of the dynamic mutation, the size of this allele could not be correlated significantly to the extent of transmission instability. This may be due to the fact that the range of expansions observed in 44 parents was confined between 34 and 54 repeats. Within this sample, no individual or sub-range of repeat length could significantly show a differential propensity for instability (Fig. 4).

DISCUSSION

The mutation causing SCA7 comprises a transcribed translated CAG repeat; this is a common motif among the 'type I' [as

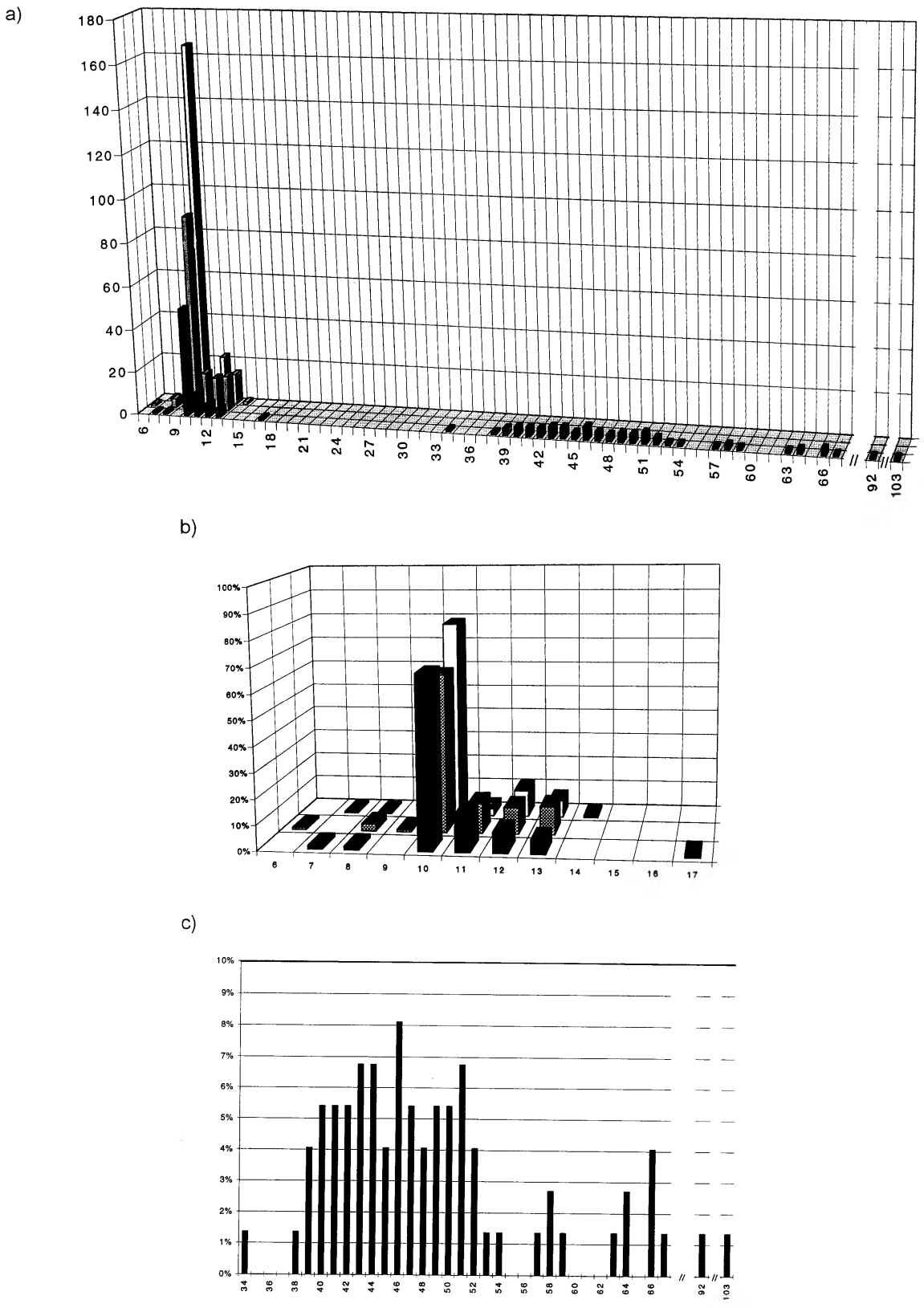

Figure 1. Repeat length of SCA7 alleles. The x-axis is the number of CAG repeats in all three graphs. The y-axis is total chromosomes in (a) and allele frequency in (b) and (c). (**a**) The filled columns in the first row represent 146 alleles from affected and pre-symptomatic individuals considered to be at risk by genotype. The shaded columns in the second row represent 144 alleles from unaffected family members. The distribution of the 210 unexpanded alleles from unaffected unrelated individuals is shown in the open columns in the rear. (**b**) Similar frequency distribution of unexpanded alleles in three populations. Columns are as described above. (**c**) Expanded SCA7 alleles ($n = 74$) show a wide distribution from 34 to 103 repeats; no single allele represents >9% of the total.

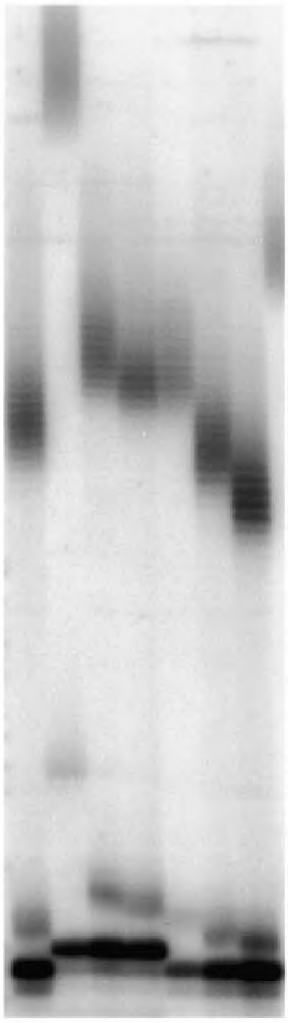

Figure 2. The autoradiogram of amplified products spanning the CAG repeat in SCA7 shows sharp bands and no heterogeneity in unexpanded alleles (bottom) but a range of products in expanded alleles (top). The repeat length of expanded alleles was determined by the darkest (usually central) band, which is flanked by 1–3 larger and smaller PCR products, which are usually lighter. Details of the PCR amplification are given in the text.

categorized in (19)] trinucleotide expansion diseases which include HD, SCA1, SCA2, SCA3, SCA6, DRPLA and SBMA. Other notable similarities with these diseases are autosomal dominant inheritance, generational anticipation, paternal bias and a cell-specific neurodegenerative phenotype despite widespread gene expression.

We have analyzed the dynamic mutation in 17 SCA7 families and correlated repeat length expansion with phenotype and inheritance pattern, and typed a large sample of independent normal chromosomes for comparative analysis. This study represents the largest analysis of SCA7 repeat length in affected and unaffected individuals to date. The observed difference between normal and expanded alleles is 17 repeats, a relatively large differential in the type I group. This striking contrast between normal and expanded alleles raises the possibility that predictive assays for disease potential in at-risk individuals might be relatively straightforward. Nevertheless, despite a correlation between quantifiable repeat length and age of onset and disease course, a genotypic assay based on SCA7 allele size cannot provide sufficient predictive value for clinical prognosis.

Although the range of repeat expansions seen in affected individuals is similar to that seen in many of the aforementioned diseases, it is notable that the normal alleles show low polymorphism and are constrained to a narrow range. Furthermore, with >97% of the unexpanded alleles harboring between 10 and 13 CAG repeats, the normal range is centered around a shorter CAG repeat length relative to the normal ranges seen in all other polyglutamine tract diseases except for SCA6 (25). We (data not shown) and others (18) have noted that in normal SCA7 alleles, the CAG tract lacks interrupting trinucleotides. As such intervening sequence have the hypothesized role of stabilizing repeat length (26), it is paradoxical that unexpanded alleles show such low heterogeneity. Despite a lack of interrupting trinucleotide within the CAG tract, low repeat length alleles could remain quite stable, explaining the limited polymorphism seen in unexpanded alleles in our sample. However, expansions beyond the low, narrow range seen in normal chromosomes would be expected to be highly unstable within populations. Indeed, this marked instability is seen in expanded alleles, and appears to be more dramatic than in other known polyglutamine tract diseases. As instability and subsequent repeat expansion contribute to age of onset and severity of disease, selective pressures might be expected to exist to keep repeat length below a threshold where the allele becomes prone to instability. An initiating gene conversion event causing duplication of the repeat sequence (27), might precipitate the appearance of the expanded allele; it is interesting to note in this regard that the smallest expanded allele is exactly twice as large as the largest unexpanded allele (both alleles occur in the same individual, who is currently asymptomatic at the age of 48). Due to the highly unstable nature of the repeat, the subsequent morbidity and the autosomal dominant nature of the disease, this mutation would not be expected to be transmitted intact through many generations before becoming incompatible with further transmission; this is consistent with the fact that SCA7 has a lower prevalence than most of the type I diseases (1,2,14). Unequal sister chromatid exchange, polymerase slippage or changes in *cis*- (28) or *trans*-acting elements [reviewed in (29) and (30)] may also be responsible for initial and/or subsequent events conferring instability on the normal CAG repeat; more intensive sequence analysis on this large patient sample may give insight into these possibilities.

The paternal bias in transmission of expanded alleles is particularly noticeable, and extends the pattern seen in other type I dynamic mutations. What is appreciated in this large sample of SCA7 kindreds is the further observation that actual disease transmission shows a strong maternal bias. Although fathers, on average, transmit significantly larger alleles to their offspring, mothers appear to be more consistent in passing on expanded alleles and thus maintaining disease within kindreds. Although females comprise 57% of total SCA7 family members and make up 64% of individuals harboring expanded alleles, they are responsible for a statistically significant 75% of all expanded allele transmissions and >80% of disease transmissions determined by pedigree history. This pattern can also be seen in previously described SCA7 families (16,31–36). In a review of 17 other described kindreds with an autosomal dominant degenerative cerebellar/retinal phenotype where a geneology was shown, or data regarding sex of affecteds and transmitting parents

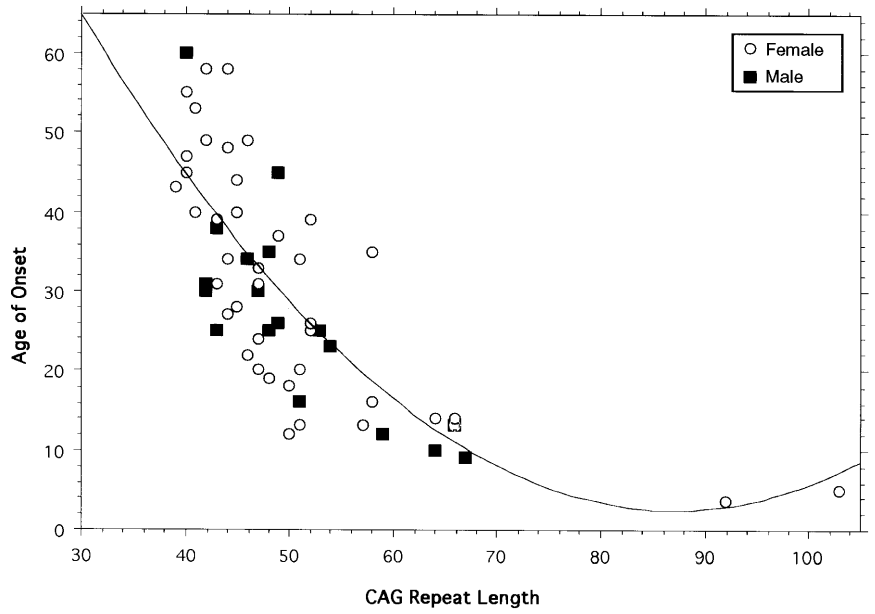

Figure 3. Quadratic regression analysis of repeat length versus age of onset gives a Pearson correlation coefficient of $r = -0.80$ $P<0.0001$. Males (squares) show a slightly, but not significantly earlier age of onset for a given repeat length.

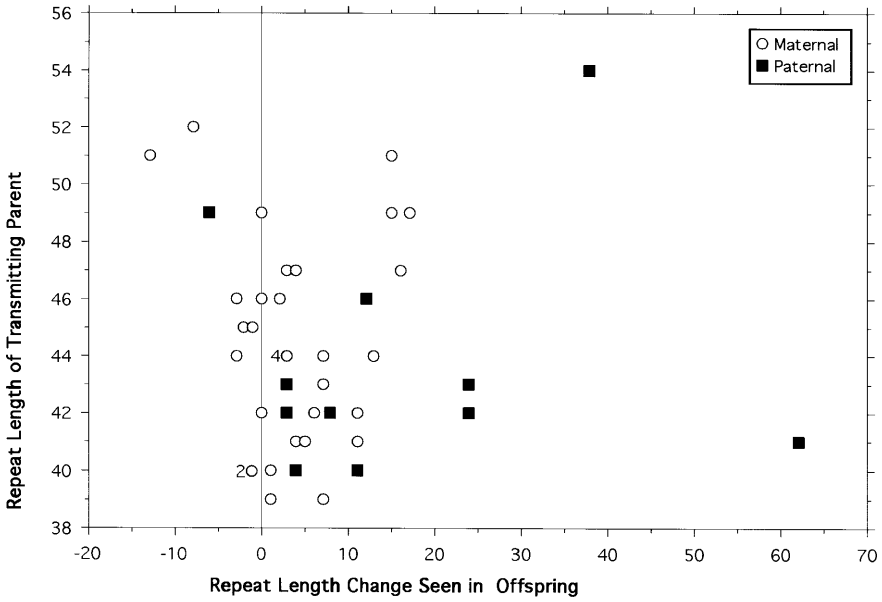

Figure 4. Parental effects on expanded allele transmission. Paternal transmission (filled squares) is responsible for more dramatic expansion changes seen in offspring, although more maternal transmissions (shaded circles) of expanded alleles occur.

could be otherwise delineated, the sex of the disease-transmitting parent was predominantly female (84 maternal transmissions versus 56 paternal). Although the significance ($P<0.05$) of this statistic in the meta-study is not as strong as that seen in our kindreds, the skew is still striking when one considers the fact that the actual ratio of affected females to males is almost exactly 50% (82 affected females versus 83 males).

Possible explanations for this female disease-transmitting bias include a potential selective disadvantage for unstable, greatly expanded paternal alleles at the gametic stage (i.e. spermatogenesis), at fertilization or *in utero*. This is entirely consistent with observations of greater allele instability in sperm populations seen in other trinucleotide repeat expansion disease (37–41). If this was the case, one would expect fathers with expanded alleles

to have significantly fewer than 0.50 of total progeny harboring expanded alleles. Where genotypes of sibships derived from affected fathers are known, this is not seen, but this sample size is quite small (6/11 children of affected fathers with expanded SCA7 alleles). It is unknown whether increased fetal wastage occurs in families where the father is affected. An alternate explanation of the observed bias takes both genetic and social factors into account: as children of affected fathers would be expected to be more severely affected with earlier age of onset, this could impact the father's decision to have more children. Indeed, although the percentage of affected males with children is approximately equivalent to that of affected females with children (83% versus 90%), sibships derived from affected fathers are smaller than those from affected mothers (mean paternal sibship size is 2.4 ± 1.7, mean maternal sibship size is 4.0 ± 3.1).

Nonetheless, the possibility that *in utero* pathology resulting from the presence of a greatly expanded paternal allele remains intriguing as there are hints that the SCA7 protein product, ataxin-7, may act as a transcription factor; as such it may play a role in development. The polyglutamine and polyproline tracts in ataxin-7 resemble those found in certain homeodomain proteins and transcription factors (42,43). Expanded ataxin-7 has been localized to the nuclear fraction using an antibody to polyglutamines (21), and a putative nuclear localization sequence in the SCA7 open reading frame suggests that the normal protein may also act within the nucleus. Study of other type I proteins offer tantalizing, albeit indefinite clues: the androgen receptor (in SBMA) is a known transcription factor, and huntingtin also carries polyglutamine/polyproline tracts (44) and is implicated as necessary for development (45–47). Further examination of ataxin-7 directed towards understanding its potential for transcriptional regulation as well as interaction with other proteins should help to elucidate its cellular role as well as its cell-specific pathogenicity when mutated.

MATERIALS AND METHODS

Subjects

The previously reported SCA7 kindreds constituted 75 family members, 38 of who were predicted to carry the SCA7 disease allele by genotype; of these 38, 25 were deemed clinically affected by exam. Fourteen additional SCA7 kindreds incorporating 70 individuals were included in the analysis. Two large and nine smaller kindreds were of European origin, one large and two smaller kindreds were African–American, one large kindred was native Peruvian and a small Liberian and small Korean kindred were included in the study. In all, 145 SCA7 family members were studied: 72 unaffected, 59 affected and 14 asymptomatic individuals believed to carry the disease allele by linkage analysis. To determine relative frequencies of normal alleles and heterogeneity of genetic polymorphism, 105 unaffected, unrelated individuals derived from the CEPH reference panel were analyzed.

Age of disease onset from patient histories was based on the time at which individuals first noted visual disturbances or problems with balance; visual changes are usually noted prior to difficulty with coordination, and Tritan axis color blindness is a reliable indicator of early symptoms (14). Clinical evaluation by neuro-ophthalmological exam was done to determine affected status. For patients where the disease course was known, duration from clinically detectable disease to debilitating disease (perambulatory incapacity and/or visual ability reduced to light/dark perception only) was categorized broadly by number of decades between states.

Subjects (or, in the case of minors, the responsible adult) signed a 'Consent for Participation' form which was approved by the Institutional Review Board for Human Research at the University of Utah School of Medicine.

DNA isolation and genotyping analysis

Anti-coagulated venous blood samples obtained from the examined individuals were used for direct DNA preparations and to establish lymphoblastoid cell lines by Epstein–Barr virus transformation as previously described (14). Genotyping using polymorphic microsatellite markers was as previously described (15).

Amplification of CAG repeats

Primer pairs 7ALT (5′-AAGGAGCGGAAAGAATGTCG-3′) and 4U716 (5′-CACGACTGTCCCAGCATCACTT-3′) were used to amplify human DNA. The 5′ termini of primer 7ALT were labeled with bacteriophage T4 polynucleotide kinase (Promega) and [γ-^{32}P]dATP (NEN) at 37°C for 30 min. Due to the repetitive nature of the DNA sequence and high GC content in and around the trinucleotide repeat, amplification by 'standard' polymerase chain reaction (PCR) was inefficient, especially for larger repeat lengths (data not shown). In order to amplify these regions more accurately, we used a combination of *Taq* DNA polymerase (for processivity) and *Pwo* DNA polymerase (for proofreading and exonuclease activity) (48). Conditions were as follows: genomic DNA (50–100 ng) in 25 µl of total reaction volume 50 mM Tris–HCl (pH 9.2) 16 mM $(NH_4)_2SO_4$, 1.75 mM $MgCl_2$, 280 µM dATP, dTTP and dCTP, 80 µM dGTP (Pharmacia), 200 µM 7-deaza-2′-dGTP (Boehringer Mannheim), 50 pM each of primers 7ALT and 4U716, 5 pM 5′-[^{32}P]7ALT, 1.5 U of *Pwo* DNA polymerase (Boehringer Mannheim), 3.5 U of *Taq* DNA polymerase (Boehringer Mannheim). Thermocycling conditions were 95°C for 5 min followed by 30 cycles of 95°C for 1 min denaturion, 56.5°C for 1 min annealing, 72°C for 1 min extension, followed by a 7 min final extension step at 72°C. After addition of 10 µl of formamide loading buffer (98% formamide, 0.01 M NaOH, 0.01% each of xylene cyanol and bromophenol blue), PCR products were denatured at 95°C for 10 min and placed on ice. Between 3 and 6 µl of product were loaded and separated by denaturing polyacrylamide gel electrophoresis (PAGE) on a 7% bis-acrylamide/acrylamide gel (Bio-Rad) 1× TBE, with 32% formamide, 6 M urea. Products were visualized either by autoradiography (Fuji X-ray film) or by PhosphorImager (Molecular Dynamics) analysis.

Sequencing analysis

Selected sequence was derived from PCR products generated as described above. Products were cycle sequenced using dideoxynucleotide terminators with SequeTherm EXCEL polymerase (Epicentre Technologies), run on PAGE and visualized by autoradiography. Normal alleles from homozygotes (CAG_{10} and CAG_{11}) as well as an expanded allele (CAG_{44}) were sequenced.

Statistical analysis

Regression coefficients, ANOVA table, mean and standard deviation data were fit using the StatView statistical analysis package version 4.5 for Apple Macintosh (Abacus Concepts, Inc.). Determination of significance for comparison of percentages was performed by χ^2 analysis, with Yates correction when necessary. Non-parametric analysis comparing means was done by the Mann–Whitney U test. The heterozygosity of unaffected individuals was determined by the equation $HET = 1 - \Sigma p_i^2$ ($i = 1...n$).

ACKNOWLEDGEMENTS

We are extremely grateful to the SCA7 families without whom this study would have been impossible. We would like to thank Alexis Brice for advance sequence information, Stephanie Zone, Grant Buetler and R. J. McKinlay Gardner for clinical coordination, and Mark Leppert, Kevin Flanigan and Ying-Hui Fu for helpful comments. This study was supported by the Howard Hughes Medical Institute, the Eccles Program in Human Molecular Biology & Genetics, NIH grant NS32711, Public Health Service research grant no. M01-RR00064 from the National Center for Research Resources, and the Utah Technology Access Section of the Utah Genome Science and Technology Center (5-P30-HG00199). L.J.P. is a Charles E. Culpeper Foundation Scholar. L.G.G. is a Howard Hughes Predoctoral Fellow in Biological Sciences.

REFERENCES

1. Harding, A. (1982) The clinical features and classification of the late onset autosomal dominant cerebellar ataxias. *Brain*, **105**, 1–28.
2. Koningsmark, B.W. and Weiner, L.P. (1970) The olivopontocerebellar atrophies: a review. *Medicine*, **49**, 227–241.
3. Yakura, H., Wakisaka, A., Fujimoto, S. and Itakura, K. (1974) Hereditary ataxia and HLA genotypes. *N. Engl. J. Med.*, **291**, 154–155.
4. Zoghbi, H.Y., Jodice, C., Sandkuijl, L.A., Kwiatkowski, T.J., McCall, A.E., Huntoon, S.A., Lulli, P. and Xenious, X. (1991) The gene for autosomal dominant spinocerebellar ataxia (SCA1) maps telomeric to the HLA complex and is closely linked to the D6S89 locus in three large kindreds. *Am. J. Hum. Genet.*, **49**, 23–30.
5. Takiyama, Y., Nishizawa, M., Tanaka, H., Kawashima, S., Sakamoto, H., Karuube, Y., Shimazaki, H., Soutome, M., Endo, K., Ohta, S., Kagawa, Y., Kanazawa, I., Mizuno, Y., Yoshida, M., Yuasa, T., Horikawa, Y., Oyanagi, K., Nagai, H., Kondo, T., Inuzuka, T., Onodera, O. and Tsuji, S. (1993) The gene for Machado–Joseph disease maps to human chromosome 14q. *Nature Genet.*, **4**, 300–304.
6. Gispert, S., Twells, R., Orozco, G., Brice, A., Weber, J., Heredero, L., Scheufler, K., Riley, B., Allotey, R., Nothers, C., Hillermann, R., Lunkes, A., Khati, C., Stevanin, G., Hernandez, A., Magarino, C., Klockgether, T., Durr, A., Chneiweiss, H., Enczmann, J., Farrall, M., Beckmann, J., Mullan, M., Wernet, P., Agid, Y., Freund, H.-J., Williamson, R., Auburger, G. and Chamberlain, S. (1993) Chromosomal assignment of the second locus for autosomal dominant cerebellar ataxia (SCA2) to chromosome 12q23–24.1. *Nature Genet.*, **4**, 295–298.
7. Ranum, L.P.W., Schut, L.J., Lundgren, J.K., Orr, H.T. and Livingston, D.M. (1994) Spinocerebellar ataxia type 5 in a family descended from the grandparents of President Lincoln maps to chromosome 11. *Nature Genet.*, **8**, 280–284.
8. Flanigan, K., Gardner, K., Alderson, K., Galster, B., Otterud, B., Leppert, M.F., Kaplan, C. and Ptacek, L.J. (1996) Autosomal dominant spinocerebellar ataxia with sensory axonal neuropathy (SCA4): clinical description and genetic localization to chromosome 16q22.1. *Am. J. Hum. Genet.*, **59**, 392–399.
9. Orr, H.T., Chung, M.-y., Banfi, S., Jr., T.J.K., Servadio, A., Beaudet, A.L., McCall, A.E., Duvick, L.A., Ranum, L.P.W. and Zoghbi, H.Y. (1993) Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type I. *Nature Genet.*, **4**, 221–226.
10. Banfi, S., Servadio, A., Chung, M.-Y., Kwiatkowski, T.J., McCall, A.E., Duvick, L.A., Shen, Y., Roth, E.J. and Orr, H.T. (1994) Identification and characterization of the gene causing type 1 spinocerebellar ataxia. *Nature Genet.*, **7**, 513–519.
11. Pulst, S.M., Nechiporuk, A., Nechiporuk, T., Gispert, S., Chen, X.N., Lopes Cendes, I., Pearlman, S., Starkman, S., Orozco Diaz, G., Lunkes, A., DeJong, P., Rouleau, G.A., Auburger, G., Korenberg, J.R., Figueroa, C. and Sahba, S. (1996) Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nature Genet.*, **14**, 269–276.
12. Imbert, G., Saudou, F., Yvert, G., Devys, D., Trottier, Y., Garnier, J.M., Weber, C., Mandel, J.L., Cancel, G., Abbas, N., Durr, A., Didierjean, O., Stevanin, G., Agid, Y. and Brice, A. (1996) Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats. *Nature Genet.*, **14**, 285–291.
13. Sanpei, K., Takano, H., Igarashi, S., Sato, T., Oyake, M., Sasaki, H., Wakisaka, A., Tashiro, K., Ishida, Y., Ikeuchi, T., Koide, R., Saito, M., Sato, A., Tanaka, T., Hanyu, S., Takiyama, Y., Nishizawa, M., Shimizu, N., Nomura, Y., Segawa, M., Iwabuchi, K., Eguchi, I., Tanaka, H., Takahashi, H. and Tsuji, S. (1996) Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT. *Nature Genet.*, **14**, 277–284.
14. Gouw, L.G., Digre, K.B., Harris, C.P., Haines, J.H. and Ptacek, L.J. (1994) Autosomal dominant cerebellar ataxia with retinal degeneration: clinical, neuropathologic and genetic analysis of a large kindred. *Neurology*, **44**, 1441–1447.
15. Gouw, L.G., Kaplan, C.D., Haines, J.H., Digre, K.B., Rutledge, S.L., Matilla, A., Leppert, M., Zoghbi, H.Y. and Ptacek, L.J. (1995) Retinal degeneration characterizes a spinocerebellar ataxia mapping to chromosome 3p. *Nature Genet.*, **10**, 89–93.
16. Benomar, A., Krols, L., Stevanin, G., Cancel, G., LeGuern, E., David, G., Ouhabi, H., Martin, J.J., Durr, A., Zaim, A. *et al.* (1995) The gene for autosomal dominant cerebellar ataxia with pigmentary macular dystrophy maps to chromosome 3p12–p21.1. *Nature Genet.*, **10**, 84–88.
17. Holmberg, M., Johansson, J., Forsgren, L., Heijbel, J., Sandgren, O and Holmgren, G. (1995) Localization of autosomal dominant cerebellar ataxia with retinal degeneration and anticipation to chromosome 3p12–21.1. *Hum. Mol. Genet.*, **4**, 1441–1445.
18. David, G., Abbas, N., Stevanin, G., Durr, A., Yvert, G., Cancel, G., Weber, C., Imbert, G., Saudou, F., Antoniou, E., Drabkin, H., Gemmill, R., Giuniti, P., Benomar, A., Wood, N., Ruberg, M., Agid, Y., Mandel, J.-L. and Brice, A. (1997) Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. *Nature Genet.*, **17**, 65–70.
19. Reddy, P. and Housman, D. (1997) The complex pathology of trinucleotide repeats. *Curr. Opin. Cell Biol.*, **9**, 364–372.
20. Koshy, B. and Zoghbi, H. (1997) THe CAG/polyglutamine tract diseases: gene products and molecular pathogenesis. *Brain Pathol.*, **7**, 927–942.
21. Trottier, Y., Lutz, Y., Stevanin, G., Imbert, G., Devys, D., Cancel, G., Saudou, F., Weber, C., David, G., Tora, L. *et al.* (1995) Polyglutamine expansion as a pathological epitope in Huntington's disease and four dominant cerebellar ataxias. *Nature*, **378**, 403–406.
22. Lindblad, K., Savontaus, M.L., Stevanin, G., Holmberg, M., Digre, K., Zander, C., Ehrsson, H., David, G., Benomar, A., Nikoskelainen, E., Trottier, Y., Holmgren, G., Ptacek, L.J., Anttinen, A., Brice, A. and Schalling, M. (1996) An expanded CAG repeat sequence in spinocerebellar ataxia type 7. *Genome Res.*, **6**, 965–971.
23. Kawakami, H., Maruyama, H., Nakamura, S., Kawaguchi, Y., Kakizuka, A., Doyu, M. and Sobue, G. (1995) Unique features of the CAG repeats in Machado–Joseph disease. *Nature Genet.*, **9**, 344–345.
24. Jodice, C., Malaspina, P., Persichetti, F., Novelletto, A., Spadaro, M., Giunti, P., Morocutti, C., Terrenato, L., Harding, A. and Frontali, M. (1994) Effect of trinucleotide repeat length and parental sex on phenotypic variation in spinocerebellar ataxia I. *Am. J. Hum. Genet.*, **54**, 959–965.
25. Zhuchenko, O., Bailey, J., Bonnen, P., Ashizawa, T., Stockton, D., Amos, C., Dobyns, W., Subramony, S., Zoghbi, H. and Lee, C. (1997) Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the a_{1A}-voltage-dependent calcium channel. *Nature Genet.*, **15**, 62–68.
26. Chung, M.Y., Ranum, L.P., Duvick, L.A., Servadio, A., Zoghbi, H.Y. and Orr, H.T. (1993) Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type I. *Nature Genet.*, **5**, 254–258.
27. Jeffreys, A., Tamaki, K., MacLeod, A., Monckton, D., Neil, D. and Armour, J. (1994) Complex gene conversion events in germline mutation at human minisatellites. *Nature Genet.*, **6**, 136–145.

28. Monckton, D.G., Neumann, R., Guram, T., Fretwell, N., Tamaki, K., MacLeod, A. and Jeffreys, A.J. (1994) Minisatellite mutation rate variation associated with a flanking DNA sequence polymorphism. *Nature Genet.*, **8**, 162–170.
29. Richards, R.I. and Sutherland, G.R. (1994) Simple repeat DNA is not replicated simply [news]. *Nature Genet.*, **6**, 114–116.
30. Harding, R., Boyce, A. and Clegg, J. (1992) The evolution of tandemly repetitive DNA: recombination rules. *Genetics*, **132**, 847–859.
31. Anttinen, A., Nikoskelainen, E., Marttila, R., Grenman, R., Falck, B., Aarnisalo, E. and Kalimo, H. (1986) Familial olivopontocerebellar atrophy with macular degeneration: a seperate entity among the olivopontocerebellar atrophies. *Acta. Neurol. Scand.*, **73**, 180–190.
32. Cooles, P., Michaud, R. and Best, P. (1988) A dominantly inherited progressive disease in a black family characterised by cerebellar and retinal degeneration, external ophthalmoplegia and abnormal mitochondria. *J. Neurol. Sci.*, **87**, 275–288.
33. Enevoldson, T.P., Sanders, M.D. and Harding, A.E. (1994) Autosomal dominant cerebellar ataxia with pigmentary macular dystrophy: a clinical and genetic study of eight families. *Brain*, **117**, 445–460.
34. Holmberg, M., Johansson, J., Forsgren, L., Heijbel, J., Sandgren, O. and Holmgren, G. (1995) Localization of autosomal dominant cerebellar ataxia associated with retinal degeneration and anticipation to chromosome 3p12–p21.1. *Hum. Mol. Genet.*, **4**, 1441–1445.
35. Jampel, R., Okazaki, H. and Bernstein, H. (1961) Ophthalmoplegia and retinal degeneration associated with spinocerebellar ataxia. *Arch. Ophthalmol.*, **66**, 247–259.
36. Weiner, L., Koningsmark, B., Stoll, J. and Magladery, J. (1967) Hereditary olivopontocerebellar atrophy with retinal degeneration: report of a family through six generations. *Arch. Neurol.*, **16**, 364–376.
37. Chong, S.S., McCall, A.E., Cota, J., Subramony, S.H., Orr, H.T., Hughes, M.R. and Zoghbi, H.Y. (1995) Gametic and somatic tissue-specific heterogeneity of the expanded SCA1 CAG repeat in spinocerebellar ataxia type 1. *Nature Genet.*, **10**, 344–350.
38. Watanabe, M., Abe, K., Aoki, M., Yasuo, K., Itoyama, Y., Shoji, M., Ikeda, Y., Iizuka, T., Ikeda, M., Shizuka, M., Mizushima, K. and Hirai, S. (1996) Mitotic and meiotic stability of the CAG repeat in the X-linked spinal and bulbar muscular atrophy gene. *Clin. Genet.*, **50**, 133–137.
39. MacDonald, M.E., Barnes, G., Srinidhi, J., Duyao, M.P., Ambrose, C.M., Myers, R.H., Gray, J., Conneally, P.M., Young, A., Penney, J. *et al.* (1993) Gametic but not somatic instability of CAG repeat length in Huntington's disease. *J. Med. Genet.*, **30**, 982–986.
40. Cancel, G., Durr, A., Didierjean, O., Imbert, G., Burk, K., Lezin, A., Belal, S., Benomar, A., Abada Bendib, M., Vial, C., Guimaraes, J., Chneiweiss, H., Stevanin, G., Yvert, G., Abbas, N., Saudou, F., Lebre, A.S., Yahyaoui, M., Hentati, F., Vernant, J.C., Klockgether, T., Mandel, J.L., Agid, Y. and Brice, A. (1997) Molecular and clinical correlations in spinocerebellar ataxia 2: a study of 32 families. *Hum. Mol. Genet.*, **6**, 709–715.
41. Takiyama, Y., Sakoe, K., Soutome, M., Namekawa, M., Ogawa, T., Nakano, I., Igarashi, S., Oyake, M., Tanaka, H., Tsuji, S. and Nishizawa, M. (1997) Single sperm analysis of the CAG repeats in the gene for Machado–Joseph disease (MJD1): evidence for non-Mendelian transmission of the MJD1 gene and for the effect of the intragenic CGG/GGG polymorphism on the intergenerational instability. *Hum. Mol. Genet.*, **6**, 1063–1068.
42. Perutz, M., Johnson, T., Suzuki, M. and Finch, J. (1994) Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. *Proc. Natl Acad. Sci. USA*, **91**, 5355–5358.
43. Gerber, H.P., Seipel, K., Georgiev, O., Hofferer, M., Hug, M., Rusconi, S. and Schaffner, W. (1994) Transcriptional activation modulated by homopolymeric glutamine and proline stretches. *Science*, **263**, 808–811.
44. The Huntington's Disease Collaborative Research Group (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell*, **72**, 971–983.
45. Duyao, M.P., Auerbach, A.B., Ryan, A., Persichetti, F., Barnes, G.T., McNeil, S.M., Ge, P., Vonsattel, J.P., Gusella, J.F., Joyner, A.L. *et al.* (1995) Inactivation of the mouse Huntington's disease gene homolog Hdh. *Science*, **269**, 407–410.
46. Nasir, J., Floresco, S.B., O'Kusky, J.R., Diewert, V.M., Richman, J.M., Zeisler, J., Borowski, A., Marth, J.D., Phillips, A.G. and Hayden, M.R. (1995) Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes. *Cell*, **81**, 811–823.
47. Zeitlin, S., Liu, J.P., Chapman, D.L., Papaioannou, V.E. and Efstratiadis, A. (1995) Increased apoptosis and early embryonic lethality in mice nullizygous for the Huntington's disease gene homologue. *Nature Genet.*, **11**, 155–163.
48. Maruyama, H., Kawakami, H. and Nakamura, S. (1996) Reevaluation of the exact CAG repeat length in hereditary cerebellar ataxias using highly denaturing conditions and long PCR. *Hum. Genet.*, **97**, 591–595.